IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

* * * * *

Methods of Forming Silicon Nitride, Methods of Forming Transistor Devices, and Transistor Devices

* * * * *

INVENTOR

John T. Moore

ATTORNEY'S DOCKET NO. MI22-1374

Methods of Forming Silicon Nitride, Methods of Forming Transistor Devices, and Transistor Devices

TECHNICAL FIELD

The invention pertains to methods of forming silicon nitride, and particularly pertains to methods of forming silicon nitride over silicon-oxide-comprising materials. The invention also pertains to methods of forming transistor devices, and further pertains to transistor device structures.

BACKGROUND OF THE INVENTION

There are numerous semiconductor processing applications in which it is desired to form a silicon-nitride-comprising layer over a silicon-oxide-comprising layer. For instance, it can be desired to form transistor devices having silicon nitride and silicon oxide as dielectric materials between a conductive gate and a channel region. A difficulty in forming silicon nitride over silicon oxide is described with reference to Figs. 1 and 2.

Referring initially to Fig. 1, a top view of a fragment of a semiconductor structure is shown. Fragment 10 comprises an oxide surface 12 upon which nitride 14 is to be formed. Oxide surface 12 can comprise, for example, silicon dioxide; and nitride 14 can comprise, for example, silicon nitride (Si_3N_4). The silicon nitride can be deposited by,

for example, chemical vapor deposition. A problem is that the silicon nitride does not deposit readily on silicon dioxide, because there are only a few bonds available for bonding of nitrogen to silicon in silicon dioxide. Accordingly, the silicon nitride forms in small localized islands. The islands grow, and eventually merge to form a silicon nitride surface 14 that entirely covers silicon oxide 12. Such silicon nitride surface is shown in Fig. 2.

The silicon nitride material that ultimately forms the surface of Fig. 2 will typically be from about 30 to 35Å thick, and will frequently be non-uniform in thickness as it was formed from the merger of relatively thick islands (frequently the islands are about 28Å thick when they merge) so that the portions where edges of the islands merged are thinner than portions corresponding to centers of the islands. Further, it can be difficult to control the overall thickness of nitride material 14, as it is difficult to control how thick the islands of Fig. 1 will be when they finally merge.

It would be desirable to develop new methods of forming silicon nitride which overcome some or all of the above-discussed problems.

SUMMARY OF THE INVENTION

In one aspect, the invention encompasses a method of forming silicon nitride on a silicon-oxide-comprising material. The silicon-oxide-comprising material is exposed to activated nitrogen species from a

nitrogen-containing plasma to introduce nitrogen into an upper portion of the material. The nitrogen is thermally annealed within the material to bond at least some of the nitrogen to silicon proximate the nitrogen. After the annealing, silicon nitride is chemical vapor deposited on the nitrogen-containing upper portion of the material.

In another aspect, the invention encompasses a method of forming a transistor device. A silicon-oxide-comprising layer is formed over a substrate. The silicon-oxide-comprising layer is exposed to activated nitrogen from a nitrogen-containing plasma to introduce nitrogen into an upper portion of the layer. The nitrogen is thermally annealed within the layer to bond at least some of the nitrogen to silicon proximate the nitrogen. After the annealing, silicon nitride is chemical vapor deposited on the nitrogen-containing upper portion of the layer. At least one conductive gate layer is formed over the silicon nitride, and defines a gate layer. A pair of source/drain regions are formed proximate the gate layer and gatedly connected to one another through a channel region that is beneath the gate layer.

In yet another aspect, the invention encompasses transistor device structures.

BRIEF DESCRIPTION OF THE DRAWINGS

Preferred embodiments of the invention are described below with reference to the following accompanying drawings.

Fig. 1 is a diagrammatic, fragmentary, top view of a semiconductor wafer fragment at a preliminary step of a prior art processing sequence.

Fig. 2 is a view of the Fig. 1 wafer fragment shown at a prior art processing step subsequent to that of Fig. 1.

Fig. 3 is a diagrammatic, cross-sectional view of a semiconductor wafer fragment at a preliminary step of a method of the present invention.

Fig. 4 is a view of the Fig. 3 wafer fragment shown at a processing step subsequent to that of Fig. 3.

Fig. 5 is a view of the Fig. 3 wafer fragment shown at a processing step subsequent to that of Fig. 4.

Fig. 6 is a view of the Fig. 3 wafer fragment shown at a processing step subsequent to that of Fig. 5.

Fig. 7 is a view of the Fig. 3 wafer fragment shown at a processing step subsequent to that of Fig. 6.

Fig. 8 is a diagrammatic, cross-sectional view of an apparatus which can be utilized in methodology of the present invention.

Fig. 9 is a diagrammatic, cross-sectional view of another apparatus which can be utilized in methodology of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

This disclosure of the invention is submitted in furtherance of the constitutional purposes of the U.S. Patent Laws "to promote the progress of science and useful arts" (Article 1, Section 8).

A method of the present invention is described with reference to Figs. 3-9.

Referring initially to Fig. 3, a wafer fragment 20 is illustrated. Fragment 20 comprises a substrate 16 having an oxide layer 18 formed thereover. Substrate 16 can comprise, for example, a bulk semiconductive material substrate such as, for example, monocrystalline silicon lightly doped with a background p-type dopant. To aid in interpretation of the claims that follow, the terms "semiconductive substrate" and "semiconductor substrate" are defined to mean any construction comprising semiconductive material, including, but not limited to, bulk semiconductive materials such as a semiconductive wafer (either alone or in assemblies comprising other materials thereon), and semiconductive material layers (either alone or in assemblies comprising other materials). The term "substrate" refers to any supporting structure, including, but not limited to, the semiconductive substrates described above.

Oxide layer 18 can be formed over substrate 16 by, for example, chemical vapor deposition of silicon dioxide. Alternatively, oxide layer 18 can correspond to a thin layer of native oxide resulting from exposure

of substrate 16 to air or other sources of oxygen. If layer 18 corresponds to native oxide, it can correspond to a layer less than or equal to approximately 5Å thick, and comprising silicon dioxide. Layer 18 can also consist essentially of silicon dioxide, or consist of silicon dioxide. In particular embodiments of the invention, layer 18 can have a thickness of from about 1Å to about 50Å, such as, for example, a thickness of about 40Å.

Fig. 4 illustrates a nitrogen-comprising region 22 being formed on and/or within oxide layer 18. Nitrogen-comprising region 22 is a thin layer formed by exposing silicon-oxide comprising material 18 to an activated nitrogen species formed from a nitrogen-containing plasma, with the term "activated" indicating that the nitrogen species is different than the form of nitrogen fed to the plasma. An activated nitrogen species can comprise, for example, a nitrogen ion or a nitrogen atom in an energy state higher than its ground state.

An exemplary method of providing nitrogen within oxide layer 14 is to expose layer 18 to a nitrogen-containing plasma and thereby introduce nitrogen into layer 18. Such introduction forms nitrogen-enriched upper region 22 of layer 18 and leaves a non-nitrogen-enriched region beneath the nitrogen-enriched upper region. The nitrogen-enriched upper portion of oxide layer 18 can be formed within the top 2Å of layer 18. If oxide layer 18 is less than or equal to about 5Å thick, nitrogen-enriched upper portion 22 preferably does not extend

below the top half of layer 18. If oxide layer 18 is at least 10Å thick, nitride-enriched portion 22 is preferably at least about 5Å above substrate 16, and more preferably only within the top 5Å of layer 18. It can be advantageous to keep the nitrogen above an upper surface of substrate 16 to avoid having the nitrogen migrate into upper surface 16 and disrupt operation formed in and/or over a surface of substrate 16. It is to be understood, however, that the invention encompasses embodiments wherein nitrogen migration into a surface of substrate 16 is acceptable, and in such embodiments the nitrogen-enriched portion of layer 18 can extend throughout an entirety of layer 18.

The nitrogen-containing plasma can be formed from, for example, N_2, NH_3 and/or N_2O. The plasma can be predominantly composed of nitrogen-containing species, consist essentially of nitrogen-containing species, or consist entirely of nitrogen-containing species. In exemplary embodiments, layer 18 is maintained at a temperature of less than or equal to 200°C during the exposure to the nitrogen-containing plasma. Such can alleviate diffusion of nitrogen into a lower half of oxide layer 18. Particular exemplary temperatures can be from 50°C to 200°C, with a suitable temperature being about 65°C. The nitrogen-containing plasma can be maintained with a power of from about 500 watts to about 5,000 watts during exposure of layer 18 to the plasma, and in particular embodiments can be maintained with a power of from about 500 watts to about 3,000 watts during the exposing. A pressure within

a reaction chamber comprising the plasma and oxide layer 18 can be less than 3 Torr. The time of exposure of layer 14 to the nitrogen-containing plasma is preferably for a time of less than or equal to about 1 minute, and in particular embodiments can be for a time of from about 3 seconds to about 1 minute. An exemplary process utilizes an exposure time of from about 10 seconds to about 15 seconds.

Figs. 8 and 9 illustrate exemplary apparatusses which can be utilized for forming nitrogen-comprising region 22.

Referring initially to Fig. 8, a remote plasma nitridization apparatus 200 is illustrated. Apparatus 200 comprises a plasma chamber 202 and a reaction chamber 204. Reaction chamber 204 comprises a substrate holder 206, and substrate 16 is supported within chamber 204 by holder 206. Preferably, holder 206 is configured to rotate substrate 16 during exposure of substrate 16 to activated nitrogen species. Such activated nitrogen species are formed within plasma chamber 202 by, for example, exposing N_2 and/or other nitrogen-containing materials (such as N_2O or NH_3) to plasma conditions. Exemplary plasma conditions comprise utilization of a microwave plasma generator at a power of from about 1,500 watts to about 3,000 watts, and utilizing a pressure within chamber 202 of less than 3 Torr. The plasma of chamber 202 forms activated nitrogen species which migrate along a passageway 208 into chamber 204 whereupon the species can

form nitrogen-comprising layer 22 (Fig. 4) over and/or within oxide 18 (Fig. 4).

An arrow is shown within passageway 208 to indicate migration of plasma activated nitrogen species through passageway 208. Preferably, passageway 208 is of sufficient length so that plasma 202 is at least about 12 inches from substrate 16. Such can enable highly activated nitrogen species formed within a plasma to relax prior to interaction with substrate 16, which can limit penetration of the nitrogen species into substrate 16 relative to an amount of penetration which would occur with more highly activated species. In order to further limit penetration of nitrogen species into substrate 16, substrate 16 is preferably not biased relative to the plasma within chamber 202.

Suitable operating conditions for forming a nitrogen-comprising plasma over substrate 16 can include maintaining a temperature of substrate 16 at from about 550°C to about 1,000°C, rotating the wafer comprising substrate 16 at about 90 rotations per minute (RPM), maintaining a pressure within chambers 202 and 204 of from about 0.8 Torr to about 2.8 Torr, and exposing the wafer to the nitridization conditions for from about one minute to about five minutes.

Referring next to Fig. 9, an alternative apparatus which can be utilized for forming nitrogen over and within oxide layer 18 (Fig. 4) is illustrated as apparatus 220. Apparatus 220 can be referred to as a high density plasma remote plasma nitridization (HDP-RPN) apparatus,

or simply as a plasma nitridization (PN) apparatus. Apparatus 220 comprises a reaction chamber 222 having a wafer holder 224 therein. Wafer 16 is supported on holder 224. A plasma 226 is formed above substrate 16, and preferably is maintained a distance "X" from substrate 16, with distance "X" corresponding to at least about four inches. Nitrogen is introduced into plasma 226 in the form of, for example, N_2, and activated nitrogen species are formed from the nitrogen. Suitable processing parameters for utilization of the apparatus of Fig. 9 include a wafer temperature of from 0°C to 400°C, no rotation of the substrate 16, a pressure within chamber 222 of from about 5 mTorr to about 15 mTorr (preferably of from about 5 mTorr to about 10 mTorr), and an exposure time of substrate 16 to activated nitrogen species within chamber 222 of from about 5 seconds to about 30 seconds.

The plasma treatment described above with reference to Figs. 8 and 9 preferably forms a minimum of about 5% nitrogen within the upper surface silicon dioxide of layer 18. The nitrogen formed within such upper surface is then thermally annealed to bond at least some of the nitrogen, and preferably all of the nitrogen, to silicon proximate the nitrogen. The thermal bonding of the nitrogen alleviates diffusion of the nitrogen through layer 18 (Fig. 4) and into substrate 16 (Fig. 4). The annealing can comprise thermal processing at a temperature of less than 1,100°C for a time of at least 3 seconds, and can comprise, for example,

a temperature of 700°C for at time of about 30 seconds, or 1,050°C for a time of about 5 seconds. Alternatively, the annealing can comprise rapid thermal processing (RTP) utilizing a ramp rate of at least 50°C/second to a temperature of less than 1,000°C, with such temperature being maintained for at least about 30 seconds. Suitable processing can include a temperature of about 900°C for at time of about 60 seconds.

After the thermal annealing of the nitrogen, the nitrogen of the nitrogen-enriched upper portion of oxide layer 18 can be entirely within the top 5Å of layer 18. More specifically, if oxide layer 18 is at least 10Å thick, the nitride-enriched portion is preferably at least about 5Å above substrate 16, and more preferably only within the top 5Å of layer 18 after the thermal annealing of the nitrogen. It can be advantageous to keep the nitrogen above an upper surface of substrate 16 to avoid having the nitrogen migrate into upper surface 16 and disrupt operation of a channel region ultimately formed in substrate 16. It is to be understood, however, that the invention encompasses embodiments wherein nitrogen migration into a surface of substrate 16 is acceptable, and in such embodiments the nitrogen-enriched portion of layer 18 can extend throughout an entirety of layer 18 after the thermal annealing.

Referring to Fig. 5, a layer 24 of silicon nitride is chemical vapor deposited over nitrogen-enriched portion 22 of silicon dioxide layer 18. An exemplary process for the chemical vapor deposition of silicon nitride

includes utilization of NH_3 and dichlorosilane as silicon nitride precursors, a temperature of from about 600°C to about 800°C, a pressure of from about 0.1 Torr to about 2 Torr, and a time of from about 10 minutes to about 100 minutes. Silicon-enriched portion 22 provides a surface upon which silicon nitride (for example, Si_3N_4) can be uniformly and controllably deposited. Layer 24 can thus be deposited to have a substantially planar upper surface, and accordingly a uniform thickness throughout. Further, the final thickness of layer 24 can be controllably determined. Preferably, layer 24 is formed to a thickness of from about 30Å to about 50Å, and in an exemplary embodiment is formed to a thickness of about 40Å over a 5Å thick silicon dioxide layer 18. In one aspect of the present invention, layers 18, 22 and 24 are utilized as a dielectric material beneath a transistor gate. Preferably, such dielectric material will provide a drive current comparable to that provided by a conventional gate oxide layer (i.e., by a gate dielectric material that consists essentially of silicon dioxide). Conventional gate oxide layers are typically from about 35Å thick to about 50Å thick. Since silicon nitride has a higher dielectric constant than silicon dioxide, the dielectric material corresponding to layers 24, 22 and 18 will be formed thicker than a dielectric material consisting of silicon dioxide to have the same drive current as the material consisting of silicon dioxide. For instance, in an embodiment in which silicon dioxide layer 18 has a thickness of about 10Å, layer 22 has a negligible thickness (i.e., less than

about 5Å), and layer 24 has a thickness of about 40Å; the composite layer formed from materials 18, 22 and 24 can be utilized in a transistor to accomplish a drive current corresponding approximately to that of a 35Å thick dielectric material consisting of silicon dioxide.

An advantage of utilizing silicon nitride within a gate dielectric material is that the silicon nitride enables the gate dielectric material to be thicker than if the material consisted of silicon dioxide, and yet to accomplish the same drive current as a dielectric material consisting of silicon dioxide. The thicker layer will be less affected by minor variations in thickness than would a thin layer. Accordingly, utilization of a thicker dielectric material comprising silicon nitride relative to a thinner material consisting of silicon dioxide can reduce tolerances of semiconductor processing steps, and accordingly reduce instances of device failure.

Figs. 6 and 7 illustrate methodology for forming a transistor device incorporating the composite dielectric material of layers 18, 22 and 24. Referring to Fig. 6, layers 26, 28 and 30 are formed over layer 24. Layers 26, 28 and 30 can correspond to, for example, conductively doped silicon, metal silicide, and insulative dielectric materials, respectively. More specifically, layer 26 can correspond to conductively doped silicon which is doped with either n-type or p-type dopant, and can correspond to, for example, conductively doped polycrystalline silicon or conductively doped amorphous silicon. Layer 28 can correspond to, for example,

titanium silicide or tungsten silicide, and layer 30 can correspond to silicon nitride or silicon dioxide.

Referring to Fig. 7, layers 18, 22, 24, 26, 28 and 30 are patterned into a gate structure of a transistor device 50. Layers 18, 22 and 24 correspond to a dielectric material under conductive gate material 26, and a channel region is defined to be beneath conductive material 26 and in the semiconductive material of substrate 16.

Lightly doped diffusion regions 40 are formed within substrate 16 and proximate the channel region. Lightly doped diffusion regions 40 can comprise either n-type dopant or p-type depending on the type of transistor device 50 (i.e., depending on whether the transistor device 50 corresponds to a PMOS or NMOS transistor).

Sidewall spacers 42 are formed adjacent sidewalls of the gate structure, and over lightly doped diffusion regions 40. Sidewall spacers 42 can be formed by conventional methods, such as, for example, by depositing an insulative material over substrate 16 and subsequently anisotropically etching the insulative material to define spacers 42. A suitable insulative material for spacers 42 is either silicon nitride or silicon dioxide.

Heavily doped source/drain regions 44 are formed within substrate 16, and are formed to be gatedly connected by the gate comprising conductive materials 26 and 28. Source/drain regions 44 can be more heavily doped with either n-type dopant or p-type dopant,

depending on whether device 50 is to be an NMOS device or a PMOS device. Heavily doped regions 44 will typically comprise at least 1×10^{19} atoms/cm^3 of conductivity enhancing dopant.

The relative thicknesses of 18, 22, 24, 26, 28 and 30 are shown diagrammatically, and not in proportion to the thicknesses of such layers in actual devices. In particular devices, layer 18 can comprise a thickness less than or equal to about 10Å, and can, for example, comprise a thickness of less than or equal to about 5Å. Layer 22 will have a negligible thickness, and will be formed within an upper portion of layer 18. Layer 22 can, for example, comprise a thickness of from about 1Å to about 2Å. Silicon nitride layer 24 can comprise a thickness of greater than or equal to 30Å, and in particular embodiments will comprise a thickness of at least about 40Å, and can, for example, comprise a thickness of at least about 50Å.

In compliance with the statute, the invention has been described in language more or less specific as to structural and methodical features. It is to be understood, however, that the invention is not limited to the specific features shown and described, since the means herein disclosed comprise preferred forms of putting the invention into effect. The invention is, therefore, claimed in any of its forms or modifications within the proper scope of the appended claims appropriately interpreted in accordance with the doctrine of equivalents.